

国外通用型数据知识库的政策调研与分析*

刘晶晶 [1,3]　顾立平 [1]　范少萍 [2,3]
[1](中国科学院文献情报中心　北京 100190)
[2](中国科学院兰州文献情报中心　兰州 730000)
[3](中国科学院大学　北京 100049)

摘要:【**目的**】为机构知识库与数据知识库建立政策规范进行政策调研，为图书馆员推荐数据知识库进行权益分析。【**方法**】以文献综述方法，进行政策调研与分析，梳理政策要素及其内容。【**结果**】研究得到管理者的权益与义务(建立审核机制、制定数据识别标准、颁布传播使用规定)；提交者的权益(免费存储、更新元数据、设置时滞期)与义务(保证数据来源可靠、遵守知识库政策、避免知识产权争议)；使用者的权益与义务(免费使用、遵循引用规定)。【**局限**】缺乏专业型数据知识库政策调研，未来可以建立完整的政策框架。【**结论**】建立完善的数据知识库政策，在平衡各方利益关系的基础上，有利于推进科研数据开放共享。

关键词: 科研数据　科研数据知识库　政策框架　权益分析
分类号: G237.6

1　数据知识库政策是科研数据开放共享的核心环节

数据已成为科技界的"一等公民"(First Class Citizen)，过去仅仅依赖科研论文作为学术评价的方式是一种历史局限[1]。为了有效地促进科研数据开放共享，科研人员逐渐形成了较为规范的数据发布途径和数据权益贡献体系，包括：利用数据知识库存储与发布科研数据；利用数据期刊发表经过同行评议的科研数据描述符[2]；利用学术期刊发表有可重复验证的科研数据所支撑的学术论文[3]。

三者之间的具体关系如图 1 所示[4]。由此可知，科研数据知识库(Research Data Repository, RDR)并不仅仅是一个科研数据开放共享的基础设施，更是整个规范的科研数据评价体系的基石。数据知识库的研究重点不仅是信息系统建设，还涉及数据质量审核与数据贡献评估，更重要的是作为整体科研数据发布规范政策的重要组成部分。

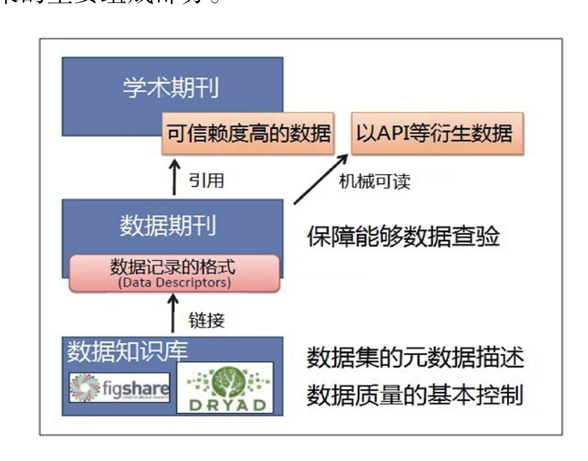


图 1　科研人员发布数据的三种途径及其关系[4]

科学研究有赖相互批评、相互成长、相互借鉴；在循环与动态的科学研究过程中，可理解的科研数据成为支持科学结论检验的重要依据[5-6]。科研数据既是科学研究的结果，也是启发进一步科学研究的宝矿[7]，

通讯作者：顾立平，ORCID：0000-0002-2284-3856，E-mail：gulp@mail.las.ac.cn。
*本文系"中国科学院文献出版领域择优支持人才"项目(项目编号:院 1434)的研究成果之一。

为了确保科研数据和相关信息的可获取、可理解和可使用，国家政府、科研资助机构、科研教育机构、科研数据中心、科研人员、图书馆员和数据知识库管理员等都发挥了重要的贡献[8]，相关权益分析逐步展开[9]，科研数据开放共享的基础环境正在形成[10]。科研数据的开放共享，不仅需要将其可靠存储，而且需要建立严谨细致的科研数据溯源管理[11]。

由于国家、基金组织、科研机构、期刊等对支撑科研成果的科研数据提出强制性存储要求和鼓励性建议，国际上建设了一批服务于科研社群的高质量数据知识库，如：Edinburgh DataShare[12]、Open Data LMU[13]、PANGAEA®[14]、Dryad[15]、figshare[16]、The Ber(li)n Digital Pantheon Project[17]等。为了帮助科研人员分辨并遴选合适的数据知识库，进而存储和重用其科研数据，数据知识库注册与目录系统应运而生，主要有OAD[18]、re3data.org[19]等，它们以全新的方式对数据知识库的相关信息进行整合、链接和呈现[20]。

中国在科研数据基础设施建设与研究方面已取得一定成果，如科研数据共享工程[21]、地球系统科研数据共享平台[22-23]、寒区旱区科学数据中心[24]、国家人口与健康科研数据共享平台[25]、国家农业科研数据共享中心[26]等，这些数据中心集中于大数据和项目数据的建设，目前已经形成一套科学数据的共享体系[27]，但对于通用型数据知识库的政策研究尚处起步阶段。

数据知识库，因其自身随着科学技术方式和环境的变化而发展，而且面对不同的学科领域、数据形态、应用层面有着不同的特点和分类[20]。根据前期调研结果[28-29]，在本文中将数据知识库分为通用型和专业型两大类，如表 1 所示：

表 1　数据知识库分类

类型	英文分类	特点	实例	面向领域
通用型数据知识库	General Research Data Repository	多学科	Dryad[15] figshare[16]	自然科学、工程技术、医学、人文和社会科学
专业型数据知识库	Special Research Data Repository	特定学科	GenBank[30]	生物医学-DNA 序列
			PubChem[31]	化学
			NOAA National Climatic Data Center [32]	气象学
			SIMBAD Astronomical Database[33]	天文学
			IQSS Dataverse Network[34]	社会科学
			Durham HepData[35]	物理学
			ORNL DAAC[36]	环境与地球科学

2　数据知识库的政策研究

2.1　研究提问

数据知识库不仅是一套信息系统，更重要的是作为科研数据开放共享的推进器。因此，传统的信息系统政策调研，注重“创建–提交–管理–使用–保存”流程管理的内容，并不适用于数据知识库调研。与之相反，如果从“用户”和“非用户”的角度进行“逆思考”，即为什么有的数据知识库得以获得广大科研人员的支持，有的数据知识库则难以被为之所用。

本研究从“利益关系人”的角度切入，包括科研资助者、科研管理者、信息服务者、出版商、科研人员等角色。围绕数据知识进行简化后，表明数据管理者(数据知识库本身的管理团队和经验团队)、数据提交者(愿意贡献科研数据的科研团队或者个人)、数据使用者(可能是同一批或者不同的科研人员)等三方角色，关系着数据知识库是否得以有效运行。

鉴于此，笔者的研究提问是：谁(第一角色)号召着谁(第二角色)存储数据并且规定谁(第三角色)如何使用数据。笔者将第一、二、三类角色所关心的问题作为本文的研究问题，总结如表 2 所示。在研究问题下的政策观测问题，即是本文分析国外通用型数据知识库政策内容的思路。

2.2　研究对象

本文对数据知识库的相关政策进行研究，以 Dryad 数据知识库有关条款[37]为主，以 figshare 为辅[38]，系统地梳理数据存储、使用、管理过程中各方的权益与义务，归纳总结国际通用型数据知识库的政策。上

表 2　数据知识库政策的分析框架

研究问题	政策观测问题
数据管理者如何进行管理?	1、如何引导数据提交者有效提交高质量的数据到数据知识库中? 2、得到数据后, 如何进行标识以便管理? 3、标识后的数据, 如何确保数据的永久可使用性? 4、得到永久可用可查的数据, 如何规范数据使用者合理使用数据知识库中的数据?
数据提交者的权利义务?	1、提交者对数据和元数据的存储、修改、增删权利是什么? 2、提交者对数据本身有怎样的自我审核义务?
数据使用者的权利义务?	1、使用者可否免费使用数据? 2、使用者在使用数据时需要注意什么?

述两个数据库中, Dryad 由一个非盈利组织管理, 为多种多样的数据类型提供一个通用的主页, 主要存储被国际数据知识库接受与出版的同行评议论文的相关科研数据集, 且不限制数据格式。figshare 目前受 Digital Science 的资助, 允许用户上传诸如图片、数据集、多媒体、论文、海报等多种形式的文档。

2.3　研究方法

本文研究设计的方法参考《前瞻导论——形塑未来与推动政策》[39]所提到"政策内容采集"到"政策建议提出"的过程, 设置观测重点。在研究过程中参考《型人导论》[40]第二章, 以实用主义方法论的四大原则判断具体内容选取与否, 并且在不同政策内容中, 以合理的方法筛选能够解释现况的最佳选择方案, 据此提炼政策要素。

3　科研数据知识库的权益管理规范

3.1　确保科研数据服务质量的审核机制

(1) 数据知识库对发布的数据进行审查和组织, 确保文档的元数据是规范、正确、可用的。Dryad 中大部分是经过同行评审的数据包[41], figshare 要求确保所提交的数据中不包含任何个人信息和医疗信息, 并且不违反英国《数据保护法》(Data Protection Act, 1998) (作为一家公司, 需要遵守注册所在地所适用的法律)[42]。

(2) 为保证提供服务的质量, 数据知识库在其声明中说明可能会进行若干审查, 包括[43]: 对包含个人的、敏感的或不合适的信息进行内容审查; 对著作权和许可声明中潜在的与 CC0 豁免规定相冲突的内容进行审查; 对内容是否符合文件格式和最低报告准则进行审查, 并在发现内容不符合条件时提醒提交者或出版商。

(3) 数据知识库要求提交者的数据文件是可被打开的、未被损坏的、没有病毒的, 而且不涉及商业利益及其可能存在的纠纷。

3.2　科研数据的标识规范

(1) 清晰说明数据来源。数据知识库的数据来源包括多种形式, 例如实验团队产生数据, 从其他数据库中获取的数据, 对他人的数据经过整理衍生而成的数据。无论是哪种数据来源, 都需要有明确说明。

(2) 以数字资源唯一标识符(Digital Object Identifier, DOI)为核心的标识规范。

例如, Dryad 中每个数据包的 DOI 通常为: "http://dx.doi.org/10.5061/dryad.[NNNN]"格式[41], 其中, "[NNNN]" 4 位数字为数据包编号, 在其后加注文件的版本信息, 格式为"/1","/2"等("/"前的数字用于表示该数据包的第几个文件)。

当新版本文件发布时, 在版本信息后加注更新信息, 如".2",".3"等。因此, 如文件信息: "http://dx.doi.org/10.5061/dryad.[NNNN].2/2.3", 则表示此文件为"http://dx.doi.org/10.5061/dryad.[NNNN]"数据包的第 2 个文件的第 2 版的第 3 次更新内容。

而 figshare 中每个数据包的 DOI 格式为: "http://dx.doi.org/10.6084/m9.figshre.[NNNNNNN]"[44], figshare 遵从 DataCite 元数据标准, 要求用户对文件添加如下信息[45]: 标题(Title)、作者列表(Authors)、类别(Categories)、标签(Tags)和一份对该文件详细的解释说明(Description)。同时为了维护研究成果的真实性和持续性, 实行版本控制, 在数据页面提供数据更新标识(Retrieved)。理论上, 数据更新后 DOI 保持不变[46]。图 2 以"All Hands to the Pump: Notes from NCCARF's 2010 International Climate Adaptation Futures Conference" [47]数据文件为例进行说明。

3.3　科研数据的长期保存

出于长期保存需要, 进行数据迁移[37]: 数据知识库在与作者签订知识产权协议的基础上, 可对数据内容的格式进行转换, 确保数据在所占容量、存储效率等方面便于传播和再利用, 保证对数据内容的日更新频率, 及时将提交者提交的内容存储到远程服务器以

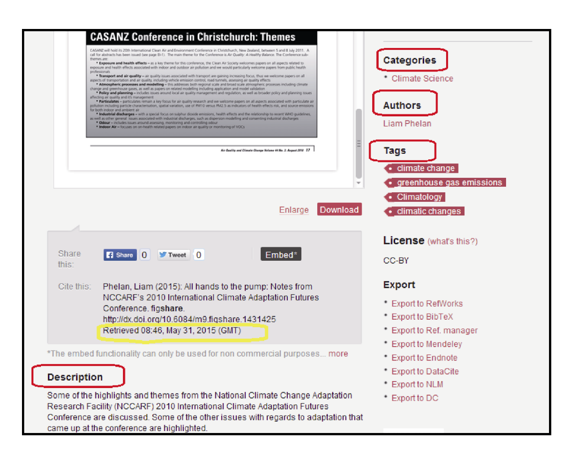


图 2　figshare 数据呈现内容

(注: 其中红色圈中是每一份数据文件必须包含的信息，黄色圈中是数据更新的呈现。)

及长期保存系统中。另外 Dryad 和 figshare 均与 CLOCKSS[48]合作保存数据内容的副本，迁移数据格式到最新版本，保证数据可无限期访问。

3.4　科研数据的传播使用规范

(1) 开放获取时期。科研数据的发布，应当考虑各方利益关系人的权益，确定合理的时滞期，数据知识库可以提供一个时间区间。如 Dryad 规定时滞期为 1-10 年[37]。

(2) 许可协议。数据知识库可以提供明确规定或者推荐许可协议(如元数据 CC0，数据 CC-BY 等)[41,49]。国际上通行的许可协议还包括：开放数据共享公共领域贡献及许可证(PDDL)等，特别强调数据库的有关条款[50]，规定了一套与数据库相关的社会规范，如开放数据共用授权社会规范等[51]。

(3) 允许发布的内容被第三方复制或索引。如 Dryad 支持与数据期刊论文和特定数据知识库(如 GenBank、DataONE、TreeBASE)数据相互关联。

(4) 撤销条款。内容发布后，数据知识库还应时刻关注其使用情况，认真记录提交者和用户提出的问题及撤销公告[37]。例如 Dryad 认为内容不合适时(包括敏感的、侵权的、不合法的资料，或者法律顾问认为其包含对 Dryad 存在风险的内容)，有权决定暂时或永久删除该内容。在对内容的元数据或其他文档记录进行修改和整合后，更新数据文件，并与出版商勘误表进行关联。(有的期刊出版商与数据知识库进行合作，共同处理数据发布的问题，例如 PLoS 和 Dryad 等多个数据知识库进行合作。)

4　科研数据提交者的权益管理规范

4.1　提交者权益

(1) 提交者享有免费存储定量数据的权益，超过该容量后，需提交附加的数据费用。为鼓励与保障数据提交工作的自觉有序进行，知识库需向数据提交者提供一定的免费存储空间。如 Dryad 提供 10GB[37]，figshare 提供 1GB 的免费容量[52]。

(2) 提交者有更新元数据的权益。Dryad 规定[37]提交者可以更新已发布数据包的元数据，以及向已发布数据包提交新的或更新的文件，且无需缴纳另外的数据发布费(仍然可能需要支付附加存储费)，对于更新的文件，同时要保留并提供获取到的所有之前发布的版本，每个文件将被赋予一个唯一标识符，且数据包的元数据将包括关于版本顺序的更新信息。在 figshare 中，提交者可以更新类目、标签和描述信息，但这些并不构成新的版本，只有对标题、作者和文件信息进行修改时才形成新的数据版本[53]。

(3) 提交者有权设置开放时滞期。在时滞期内，数据仅供项目和课题承担单位及授权范围内的用户访问和使用。时滞期过后，可向社会公众全面或分批次开放[54]。在实践中，需要注意区别开放获取时滞期、开放时期和最低保存期。

4.2　提交者义务

提交者一方面要保证数据内容的正确性、合法性，另一方面还需遵守知识库的相关规定。具体包含以下三点[37]:

(1) 数据提交者必须保证是内容的作者，或已获得内容作者的授权，可将内容发布到知识库，保证内容的正确性，不存在虚假或误导信息。

(2) 提交者在提交数据内容的同时即被认为知晓并遵守数据知识库的许可协议，并允许数据知识库依据该许可协议开放其数据，以及对内容的推广与宣传、格式转换、元数据修改、删除部分内容等操作。此外，提交者在遵守数据知识库有关内容发布的许可协议时需保证符合出版商、资助机构、用人单位等的相关准则与合同，避免可能引起的纠纷。

(3) 提交者应保证提交内容不侵犯知识产权、个人隐私权等国家的相关法律法规。

5 科研数据使用者的权益管理规范

5.1 使用者权益

数据使用者在遵循开放获取的前提下，可以免费使用科研数据。例如，在 Dryad 政策中，使用者对知识库中的数据内容享有免费下载、重加工、再利用以及交流等多项权益，但前提是需遵循知识库对数据内容的相关许可协议。科学数据的开放共享可追溯到 2010 年的开放科研数据原则"潘顿原则"(Panton Principles)，认为开放科研数据意味着允许任何用户以任何目的免费通过互联网对数据进行下载、复制、分析和再利用。同时，不受资金、法律或其他技术壁垒的限制[55]。目前，数据知识库普遍遵循潘顿原则。

5.2 使用者义务

(1) 数据知识库的引用规范。如 Dryad 要求用户采用的引用格式[41]如下：

<Creater>(<Publication Year>) Data from:<Title>. Dryad Digital Repository.<Identifier>

如图 3 所示，在该引用实例中，既包括原文的引用信息，还包括 Dryad 规定的对数据包的引用信息。

> Westbrook JW, Kitajima K, Burleigh JG, Kress WJ, Erickson DL, Wright SJ (2011) Data from: What makes a leaf tough? Patterns of correlated evolution between leaf toughness traits and demographic rates among 197 shade-tolerant woody species in a neotropical forest. Dryad Digital Repository. http://dx.doi.org/10.5061/dryad.8525

图 3 Dryad 知识库中数据内容引用格式案例[41]

同理，fighsare 在其每一份数据之下都会有"Cite this"的标识，如图 4 所示[56]：

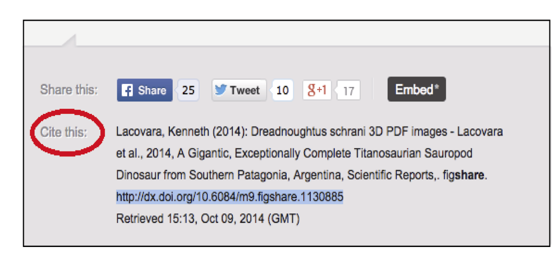


图 4 figshare 知识库中内容引用格式案例[56]

(2) 除了引用 ID 识别码外，也可以对存储科研数据的知识库网址进行引用。如 ArrayExpress 功能基因组学实验数据库要求在引用数据时包含数据的识别符和 ArrayExpress 主页网址[57]。

(3) 若干数据知识库也会推荐采用 DataCite 的引用格式或者类似格式。如 GEO(Gene Expression Omnibus) 建议提交者引用其识别符(GSExxx)，同时也建议用户引用他人的原文和该文章所对应数据记录的识别符[58]。

(4) 在使用数据知识库时，使用者必须遵守数据知识库及使用者所在当地的国家法律、法规，不得使用非法方式；不得损害其他使用者的权益，或破坏数据知识库的功能。

6 结 语

6.1 实践意义

基于以上论述，笔者提炼数据知识库的政策要点，如表 3 所示。建立数据知识库时，可以参考此表进行政策规划；图书馆员推荐科研人员存储科研数据时，也可利用此表进行权益说明；机构知识库存储与传播科研数据时，可以参照此表进行有关的政策补充条款等说明。

表 3 通用型数据知识库的政策要素

利益相关者	政策要素
数据知识库管理者	(1) 建立审核机制(元数据符合规范；保留抽查权利；数据的可使用性)； (2) 制定数据识别标准(说明数据来源；DOI 及其标示方式；新版本标示方式)； (3) 为长期保存的数据迁移程序； (4) 颁布传播使用规定(开放获取时滞期；许可协议；允许第三方利用；撤销条款)。
数据提交者	权益：(1) 免费存储定量数据；(2) 更新元数据；(3) 设置开放获取时滞期。 义务：(1) 保证数据来源可靠；(2) 遵守知识库政策；(3) 避免知识产权争议。
数据使用者	权益：免费使用数据知识库中的数据。 义务：遵守数据知识库所规范的引用方式。

6.2 研究局限

本文根据调研的部分数据知识库的有关条款，形成系统性的针对通用型数据知识库的政策框架。将研究结论进行推广时，需要注意各国法律法规以及科研生态体系的不同，并且在长短期资助经营上的考虑等。

例如，Dryad 对其存储的数据内容与元数据全部实施 CC0 许可协议，已经与数据提交者签署知识产权转让协议，并且采取一定的收费形式。国际上，有些数据知识库对存储的元数据采用 CC0 协议，而对数据本身采用 CC-BY 或者其他许可协议。

此外，数据知识库政策能否照搬到机构知识库还

是需要一定的政策调整，通用型数据知识库面向的是全球科研人员，而机构知识库面向的是具有专业属性的一批科研人员或者教职员工。无论是 Dryad 还是 figshare 等数据知识库，均保证了一定的存储空间，但超过这个空间，则要收费。一方面，数据与论文不同，可能数据集需要不断追加存储量；另一方面，数据存储涉及不同版本的数据集，因此政策中，往往规定作者可以修改元数据，但更改或者增减数据，则需要通过申请或者额外付费等。不仅在收费问题上，机构知识库的立场与通用型数据知识库不同，而且在存储内容乃至管理方式上也有所不同，这些都是值得注意的政策细节之处。

6.3　未来研究

本研究针对“通用型数据知识库”进行调研分析，对于其他“专业型数据知识库”还需进一步研究。中国科学院文献情报中心组织编译了《科研数据知识库的政策汇编》[①]报告。未来将调研我国科研人员的需求，以及我国数据知识库优秀实践案例，形成关于数据知识库政策的最佳实践汇编，提供更为具体详尽的遴选参考范例。

(致谢：感谢中国科学院计算机网络信息中心刘峰老师以及匿名评审老师的指导意见。)

参考文献：

[1] Bolikowski Ł, Houssos N, Manghi P, et al. Data as “First-class Citizens” [J/OL]. D-Lib Magazine, 2015, 21(1-2): DOI: 10.1045/january2015-guest-editorial. http://www.dlib.org/dlib/january15/01guest_editorial.print.html.

[2] 刘晶晶，顾立平．数据期刊的政策调研与分析——以 Scientific Data 为例[J]．中国科技期刊研究，2015, 26(4): 331-339. (Liu Jingjing, Ku Liping. The Policy Research and Analysis of Data Journals: Taking Scientific Data as an Example [J]. Chinese Journal of Scientific and Technical Periodical, 2015, 26(4): 331-339.)

[3] 吴蓉，顾立平，刘晶晶．国外学术期刊数据政策的调研与分析[J]．图书情报工作，2015, 59(7): 99-105. (Wu Rong, Ku Liping, Liu Jingjing. Research on the Data Policy of Academic Journals [J]. Library and Information Service, 2015, 59(7): 99-105.)

[4] 林和弘，村山泰啓．オープンサイエンスをめぐる新しい潮流(その 3)研究データ出版の動向と論文の根拠データの公開促進に向けて[J]．科学技術動向，2015(1-2): 4-9. (Hayashi K, Murayama Y. Trend of Research Data Publishing and the Supported-article Data to Public Access [J]. Science and Technology Trends – Quarterly Review, 2015(1-2): 4-9.)

[5] The Royal Society. Science as an Open Enterprise [EB/OL]. (2012-06-21). [2015-03-14]. http://royalsociety.org/policy/projects/science-public-enterprise/report/.

[6] Hey T, Tansley S, Tolle K，等．第四范式：数据密集型科学发现[M]．潘教峰，张晓林，等译．北京：科学出版社，2012. (Hey T, Tansley S, Tolle K, et al. The Fouth Paradigm: Data-Intensive Scientific Discovery [M]. Translated by Pan Jiaofeng, Zhang Xiaolin, et al. Beijing: Science Press, 2012.)

[7] Marchionini G．科研数据管理：保障数据质量，促进 iSchools 新科学研究[J]．杨冠灿，芦昆译．图书情报知识，2013(4): 4-9. (Marchionini G. Research Data Stewardship: Ensuring Data Quality to Enable New Science in iSchools[J]. Translated by Yang Guancan, Lu Kun. Documentation, Information & Knowledge, 2013(4): 4-9.)

[8] Davidson J, Jones S, Molloy L, et al. Emerging Good Practice in Managing Research Data and Research Information Within UK Universities [J]. Procedia Computer Science, 2014, 33: 215-222.

[9] 顾立平．科学数据权益分析的基本框架[J]．图书情报知识，2014(1): 34-51. (Ku Liping. Basic Framework of the Analysis on Research Data Rights [J]. Documentation, Information & Knowledge, 2014(1): 34-51.)

[10] 张晓林．开放获取、开放知识、开放创新推动开放知识服务模式——3O 会聚与研究图书馆范式再转变[J]．现代图书情报技术，2013(2): 1-10. (Zhang Xiaolin. Open Access, Open Knowledge, and Open Innovation Pushes for Open Knowledge Services——3O Convergence and a New Paradigmatic Shift for Research Libraries [J]. New Technology of Library and Information Service, 2013(2): 1-10.)

[11] Buneman P, Davidson S B. Data Provenance - The Foundation of Data Quality [EB/OL]. [2015-03-14]. http://www.sei.cmu.edu/measurement/research/upload/Davidson.pdf.

[12] Edinburgh DataShare[EB/OL].[2015-03-14]. http://datashare.is.ed.ac.uk/.

[13] Open Data LMU [EB/OL]. [2015-03-14]. http://data.ub.uni-muenchen.de/.

[14] PANGAEA® [EB/OL]. [2015-03-14]. http://www.pangaea.de/.

[15] Dryad [EB/OL]. [2015-03-14]. http://datadryad.org/.

[16] figshare [EB/OL]. [2015-03-14]. http://figshare.com/.

①该报告已在中国科学院文献情报中心机构知识库提供开放共享，编译人员：茹丽洁、戚义姣。

[17] The Ber(li)n Digital Pantheon Project [EB/OL]. [2015-03-14]. http://www.digitalpantheon.ch/repository.
[18] OAD [EB/OL]. [2015-03-14]. http://oad.simmons.edu/oadwiki/Data_repositories.
[19] re3data.org [EB/OL]. [2015-03-14]. http://re3data.org.
[20] 刘峰, 张晓林, 孔丽华. 科研数据知识库研究述评[J]. 现代图书情报技术, 2014(2): 25-31. (Liu Feng, Zhang Xiaolin, Kong Lihua. Research Review on the Research Data Repositories [J]. New Technology of Library and Information Service, 2014(2): 25-31.)
[21] 科研数据共享工程[EB/OL]. [2015-03-14]. http://baike.baidu.com/view/1047817.htm.(The Project of Research Data Sharing [EB/OL]. [2015-03-14]. http://baike.baidu.com/view/1047817.htm.)
[22] 地球系统科研数据共享平台[EB/OL]. [2015-03-14]. http://www.geodata.cn/Portal/?isCookieChecked=true. (Data Sharing Infrastructure of Earth System Science[EB/OL].[2015-03-14]. http://www.geodata.cn/Portal/?isCookieChecked=true.)
[23] 诸云强, 孙九林, 冯敏, 等. 论地学科研信息化环境[J]. 中国科学院院刊, 2013, 28(4): 501-510. (Zhu Yunqiang, Sun Jiulin, Feng Min, et al. Study on e-Science for Geosciences [J]. Bulletin of the Chinese Academy of Sciences, 2013, 28(4): 501-510.)
[24] 寒区旱区科学数据中心[EB/OL]. [2015-03-14]. http://westdc.westgis.ac.cn/. (West DC[EB/OL]. [2015-03-14]. http://westdc.westgis.ac.cn/.)
[25] 国家人口与健康科研数据共享平台[EB/OL]. [2015-03-14]. http://www.ncmi.cn/1. (National Scientific Data Sharing Platform for Population and Health [EB/OL]. [2015-03-14]. http://www.ncmi.cn/1.)
[26] 国家农业科研数据共享中心[EB/OL]. [2015-03-14]. http://www.agridata.cn/. (AgriData [EB/OL]. [2015-03-14]. http://www.agridata.cn/.)
[27] 吴立宗, 王亮绪, 南卓铜, 等. 科学数据出版现状及其体系框架[J]. 遥感技术与应用, 2013, 28(3): 383-390. (Wu Lizong, Wang Liangxu, Nan Zhuotong, et al. Scientific Data Publication: An Review and Framework [J]. Remote Sensing Technology and Application, 2013, 28(3): 383-390.)
[28] Pampel H, Vierkant P, Scholze F, 等. 呈现科研数据知识库: re3data.org 注册机制[J]. 顾立平译. 现代图书情报技术, 2014(3): 26-34. (Pampel H, Vierkant P, Scholze F, et al. Making Research Data Repositories Visible: The re3data.org Registry [J]. Translated by Ku Liping. New Technology of Library and Information Service, 2014(3): 26-34.)
[29] National Science Board. Long-Lived Digital Data Collections: Enabling Research and Education in the 21st Century [R/OL]. (2005-09-28). [2015-03-14]. http://www.nsf.gov/pubs/2005/nsb0540/nsb0540.pdf.
[30] GenBank [EB/OL]. [2015-03-14]. http://www.biosharing.org/biodbcore-000001.
[31] PubChem [EB/OL]. [2015-03-14]. http://www.biosharing.org/biodbcore-000455.
[32] NOAA National Climatic Data Center [EB/OL]. [2015-03-14]. http://service.re3data.org/repository/r3d100010286.
[33] SIMBAD Astronomical Database [EB/OL]. [2015-03-14]. http://service.re3data.org/repository/r3d100010163.
[34] IQSS Dataverse Network [EB/OL]. [2015-03-14]. http://dataverse.org/.
[35] The Durham HepData Project [EB/OL]. [2015-03-14]. http://durpdg.dur.ac.uk/abouthepdata.
[36] ORNL DAAC [EB/OL]. [2015-03-14]. http://daac.ornl.gov/.
[37] Dryad-Terms of Service [EB/OL]. (2013-08-22). [2015-03-14]. http://datadryad.org/themes/Mirage/docs/TermsOfService-Letter-2013.08.22.pdf.
[38] Using figshare and FAQs [EB/OL]. [2015-03-14]. https://figshare.zendesk.com/hc/en-us.
[39] 顾立平. 前瞻导论——形塑未来与推动政策[M]. 台北: 设计家出版社, 2013. (Ku Liping. On Foresight: Sharing Future with Forwarding Policy [M]. Taipei: Taipei Designer Press, 2013.)
[40] 顾立平. 型人导论——网络用户信息行为与差异化服务策略[M]. 科学技术文献出版社, 2013: 11-13. (Ku Liping. On Perssonas: Web User Information Behavior and Differentated Services Strategy [M]. Scientific and Technical Documentation Press, 2013: 11-13.)
[41] What Kinds of Data does Dryad Accept? [EB/OL]. [2015-03-14]. http://datadryad.org/pages/faq.
[42] How do I Publish My Data? [EB/OL]. [2015-03-14]. https://figshare.zendesk.com/hc/en-us/articles/203712033-How-do-I-publish-my-data-.
[43] Dryad Obligations,Representations&Warranties to Purchasers and Submitters [EB/OL]. [2015-03-14]. http://datadryad.org/themes/Mirage/docs/TermsOfService-Letter-2013.08.22.pdf.
[44] figshare-Browse Data[EB/OL]. [2015-03-14]. http://figshare.com/articles/browse#thumb.
[45] What Metadata does Figshare Assign to Files? [EB/OL]. [2015-03-14]. https://figshare.zendesk.com/hc/en-us/articles/203979333-What-metadata-does-figshare-assign-to-files-.
[46] How do I Upload a New Version of My Data? [EB/OL]. [2015-03-14]. https://figshare.zendesk.com/hc/en-us/articles/203727056-How-do-I-upload-a-new-version-of-my-data-.
[47] All Hands to the Pump: Notes from NCCARF's 2010

International Climate Adaptation Futures Conference [EB/OL]. [2015-05-31]. http://figshare.com/articles/All_hands_to_the_pump_Notes_from_NCCARF_s_2010_International_Climate_Adaptation_Futures_Conference/1431425.
[48] The CLOCKSS Archive [EB/OL]. [2015-03-14]. http://www.clockss.org/clockss/Home.
[49] Why does figshare Use CC Licenses? [EB/OL]. [2015-03-14]. https://figshare.zendesk.com/hc/en-us/articles/201953883-Why-does-figshare-use-CC-licenses-.
[50] Open Data Commons Public Domain Dedication and License (PDDL) [EB/OL]. [2015-03-14]. http://opendatacommons.org/licenses/pddl/.
[51] ODC Attribution-Sharealike Community Norms [EB/OL]. [2015-03-14]. http://opendatacommons.org/norms/odc-by-sa/.
[52] figshare Pricing [EB/OL]. [2015-03-14]. http://figshare.com/pricing.
[53] Does figshare Support Version Control [EB/OL]. [2015-03-14]. https://figshare.zendesk.com/hc/en-us/articles/203922158-Does-figshare-support-version-control-.
[54] 司莉, 李月婷. 国外科学数据保存政策的分析与启示[J]. 信息资源管理学报, 2014, 4(2): 40-50. (Si Li, Li Yueting. Analysis and Revelation on Research Data Preservation Policy Abroad [J]. Journal of Information Resources Management, 2014, 4(2): 40-50.)
[55] Panton Principles-Principles for Open Data in Science [EB/OL]. [2015-03-14]. http://pantonprinciples.org/.
[56] Is the Research I Put on figshare Citable / Do I Get a DOI? [EB/OL]. [2015-03-14]. https://figshare.zendesk.com/hc/en-us/articles/201954103-Is-the-research-I-put-on-figshare-citable-Do-I-get-a-DOI-.
[57] ArrayExpress-Submitting Data to ArrayExpress (general) [EB/OL]. [2015-03-14]. http://www.ebi.ac.uk/arrayexpress/help/faq.html#cite.
[58] GEO-Citing and Linking to the GEO Database [EB/OL]. [2015-03-14]. http://www.ncbi.nlm.nih.gov/geo/info/linking.html.

作者贡献声明:

刘晶晶: 论文修改, 信息补充, 案例采集和分析, 参考文献整理;
顾立平: 研究设计, 查找信息源, 政策分析, 论文最终版本修订;
范少萍: 论文起草, 提出研究问题与框架。

收稿日期: 2015-03-23
收修改稿日期: 2015-06-15

The Policy Research and Analysis of General Research Data Repository

Liu Jingjing[1,3] Ku Liping[1] Fan Shaoping[2,3]
[1](National Science Library, Chinese Academy of Sciences, Beijing 100190, China)
[2](Lanzhou Library, Chinese Academy of Sciences, Lanzhou 730000, China)
[3](University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: [**Objective**] Research and analyze the policy for establishing specification of Institutional Repository and Research Data Repository; explore related rights and obligations for recommending the data repository to librarians. [**Methods**] Summarize and refine policy elements through the literature review and Internet research. [**Results**] The results include rights and obligations of managers (establishing audit mechanism, making data identification standards, issuing regulations of spreading and using), rights of submitters (free storage, update metadata, set the embargo) and its obligations (ensuring reliable data source, abiding by the policy of data repository, avoiding the intellectual property rights disputes), rights and obligations of users (free of charge, follow the reference rules). [**Limitations**] Lack the policy research on Special Research Data Repository, thus the future study can establish a complete policy framework. [**Conclusions**] Establishing a complete policy of Research Data Repository, which can balance the interests of all parties and then promote the research data sharing.

Keywords: Research data Research Data Repository Policy framework Analyze of rights and obligations